Constructive Alignment in the Age of AI*

Claus Brabrand
Center for Computing Education Research (CCER)
IT University of Copenhagen (ITU)
Copenhagen, Denmark
brabrand@itu.dk

Paul Denny
School of Computer Science
University of Auckland
Auckland, New Zealand
paul@cs.auckland.ac.nz

Abstract

Generative AI has been a disruptive force in higher education. Responses have ranged from enthusiastic adoption to outright prohibition. We argue that many of the challenges currently attributed to AI are not fundamentally new: AI is not the *cause*; AI is the *catalyst* that exposes and intensifies longstanding systemic misalignments in educational design. Constructive Alignment emphasizes that quality learning depends on alignment between *intended learning outcomes* (ILOs), *assessment tasks* (ATs), and *teaching/learning activities* (TLAs): what competences students are expected to learn, how that learning is *incentivized* (through assessment), and how learning is *supported* (via teaching/learning activities).

Building from this foundation, we propose an extension to Constructive Alignment for the AI era that introduces a new dimension: *AI-mode* (of learning). We distinguish between the following AI-modes of learning: *AI-off* (learning *without* AI), *AI-on* (learning *with* AI), and *on-AI* (learning *about* AI). We argue that each AI-mode requires corresponding intended learning outcomes (ILOs), assessment tasks (ATs), and teaching/learning activities (TLAs), all aligned with that mode. We present three design principles for implementing alignment in the age of AI: (1) formulate the intended competence of each intended learning outcome (ILO) and designate its *AI-mode*; (2) align assessment tasks (ATs) with the designated AI-mode in order to *incentivize* learning within that AI-mode; and (3) align teaching/learning activities (TLAs) with the designated AI-mode so as to *support* learning within that AI-mode.

Our analysis reframes current debates about AI in education as questions of systemic alignment rather than technology. The resulting framework provides a stabilizing, theory-driven foundation for *teaching for quality learning in the age of AI.*

1 Introduction: The AI-*in*-Education Crisis

The emergence of Generative AI has transformed higher education, giving rise to responses ranging from enthusiastic adoption to outright prohibition. Some educators emphasize opportunities for personalized tutoring, rapid feedback, increased productivity, and new forms of learning. Others worry about academic integrity, (in)validity of assessment, over-reliance on AI, and the erosion of fundamental competences. Amid these competing perspectives, many educators and educational institutions are left bewildered, wondering what role AI should play in their teaching, learning, and assessment.

These questions are important, but they are often framed too narrowly. Discussions frequently become polarized, focusing on whether AI should be allowed or prohibited, whether particular assessments remain valid, and whether students are "cheating" when using AI. While such questions matter, we argue that they are symptoms of a deeper issue. The central challenge is one of *alignment.*

What do we intend students to learn? Under what conditions should they demonstrate that learning? What kinds of teaching/learning activities provide appropriate opportunities to develop the intended competence? These questions sit at the heart of Constructive Alignment [5], one of the most influential theories of educational design in higher education. Generative AI has not changed the importance of these questions, but it has made them considerably more difficult to ignore.

A common narrative is that AI has created entirely new educational problems. We argue instead that many of the challenges currently attributed to AI are not fundamentally new. Long before AI, students could seek assistance from friends, spouses, siblings, tutors, online resources, or previous reports. Such forms of "cognitive outsourcing" varied in availability and quality, but they nevertheless created opportunities for students to complete assessed work without fully engaging in the intended learning process. These opportunities often required social capital, financial resources, effort, and sometimes luck, limiting their scale and visibility.

Generative AI did not invent cognitive outsourcing. What it changed was its economics. Tasks that previously required access to knowledgeable people, paid tutoring, or substantial effort can now be outsourced instantly, at negligible cost, and at virtually unlimited scale. We therefore argue that AI is not the *cause*; AI is the *catalyst.* It exposes and intensifies longstanding systemic educational misalignments between intended learning outcomes (ILOs), assessment tasks (ATs), and teaching/learning activities (TLAs). Misalignment existed before AI. After the proliferation of AI, misalignment intensified and became impossible to ignore.

This observation motivates the central thesis of this paper. AI does not break Constructive Alignment. Rather, AI acts as a "stress test" of Constructive Alignment, exposing longstanding misalignments that could previously remain limited in scale and therefore largely unnoticed. The appropriate response is therefore not to abandon existing educational theory, but to apply it more rigorously than ever before.

Drawing on Constructive Alignment, we propose an extension that introduces a new dimension: *AI-mode* (of learning). We explicitly distinguish between three AI-modes of learning:

- **AI-off** (learning *without* AI);
- **AI-on** (learning *with* AI); *and*
- **on-AI** (learning *about* AI).

We argue that each AI-mode requires corresponding intended learning outcomes (ILOs), assessment tasks (ATs), and teaching/learning activities (TLAs), all aligned with that mode. This reframes the discussion from a simplistic binary question, *'Should AI be allowed?,'* to a higher-level systemic educational design question:

*EdArXiv preprint, 1 July 2026. https://doi.org/10.35542/osf.io/m9yfk_v1

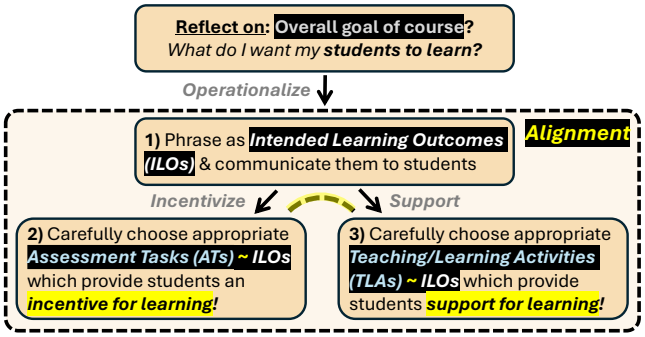

Figure 1: The Theory of Constructive Alignment, emphasizing the (constructive) alignment implementation process.

> *"**What role should AI play** in intended learning outcomes (ILOs), assessment tasks (ATs), and teaching/learning activities (TLAs)?"*

Contributions. We make three main contributions:

(1) **Relevance of Constructive Alignment:** We *reassert* Constructive Alignment as a stabilizing theoretical framework for navigating the age of AI;
(2) **Extension through AI-mode:** We *extend* Constructive Alignment by introducing *AI-mode* as an explicit design dimension across intended learning outcomes, assessment tasks, and teaching/learning activities; *and*
(3) **Diagnosis of AI-mode Misalignment:** We *present* a practical framework for the *diagnosis* and *resolution* of what we term *AI-mode misalignment*.

The remainder of the paper is organized as follows. Section 2 revisits the principles of Constructive Alignment and introduces the AI-mode dimension. Sections 3–5 examine how each of *intended learning outcomes* (ILOs), *assessment tasks* (ATs), and *teaching/learning activities* (TLAs) can be designed for *AI-off*, *AI-on*, and *on-AI* modes of learning. We conclude by discussing the implications of the framework for teaching for quality learning in the age of AI.

2 Constructive Alignment

To understand why generative AI has disrupted higher education, we first revisit the theory of Constructive Alignment [5]. We argue that AI does not invalidate Constructive Alignment; rather, it exposes long-standing misalignments that could previously remain limited in scale and therefore largely unnoticed. Consequently, the appropriate response to the current AI-*in*-Education crisis is not to abandon Constructive Alignment, but to adapt it to an AI context and apply it even more rigorously.

This section first reviews Constructive Alignment in its traditional pre-AI form (Section 2.1), before the emergence and widespread adoption of generative AI. We then show how AI acts as a catalyst that exposes and intensifies long-standing educational misalignments, motivating the introduction of an additional design dimension—AI-mode—that enables Constructive Alignment to address contemporary AI-related challenges (Section 2.2).

2.1 Constructive Alignment before AI

The Theory of Constructive Alignment is one of the most influential theories of educational design in higher education [5, 9]. At its core, it revolves around three deceptively simple questions:

(1) *What do we want students to learn?*
(2) *How do we incentivize it?*
(3) *How do we support it?*

In Constructive Alignment, these questions are operationalized through Intended Learning Outcomes (ILOs), Assessment Tasks (ATs), and Teaching/Learning Activities (TLAs), respectively.

Implementation Process. Figure 1 illustrates the corresponding alignment implementation process. The process begins with a reflection on the overall purpose of the course and the question: *What do I want my students to learn?* This overall educational goal is then operationalized through a sequence of three design steps.

(1) Intended Learning Outcomes (ILOs). The first step is to formulate explicit Intended Learning Outcomes (ILOs) and communicate them to students. ILOs specify the competences students are expected to be able to *do* upon completion of the course. They make the intended learning explicit, provide the foundation on which all subsequent educational design decisions are based, and give students a roadmap for tracking their own learning progress.

(2) Assessment Tasks (ATs). The second step is to design assessment tasks aligned with the intended learning outcomes (ATs ~ ILOs), requiring students to *demonstrate* the competences specified by the ILOs. Assessment serves a dual purpose. It *measures* learning, but it also *incentivizes* learning. Students naturally direct their efforts towards what is assessed—a phenomenon often referred to as the *backwash effect* [5]. Constructive Alignment therefore requires assessment tasks to validly measure and reward the competences specified by the intended learning outcomes.

(3) Teaching/Learning Activities (TLAs). The third step is to design teaching/learning activities aligned with the intended learning outcomes (TLAs ~ ILOs), allowing students to *develop* and *practice* the competences specified by the ILOs. While assessment tasks *incentivize* learning, teaching/learning activities *support* learning by enabling students to develop the intended competences through guided practice.

ATs ~ TLAs. Since both assessment tasks and teaching/learning activities are aligned with the intended learning outcomes, they are necessarily aligned with one another.[1] Although this is sometimes criticized as *"teaching to the test"*, it is entirely appropriate when the assessment tasks themselves are validly aligned with the intended learning outcomes. In that case, teaching to the test simply means teaching towards the intended learning.

Perceived Alignment. Importantly, alignment is not merely a matter of educational design. It must also be communicated convincingly to students so that they perceive the alignment. Otherwise, the *backwash effect* may dominate: If students believe that one set of competences is rewarded while another is merely advocated, assessment rather than intended outcomes will shape their learning behavior [5]. Constructive Alignment therefore seeks to align not only outcomes, assessment, and activities, but also educator

[1] Logically: $((\text{ILOs} \sim \text{ATs}) \wedge (\text{ILOs} \sim \text{TLAs})) \Rightarrow (\text{ATs} \sim \text{TLAs})$.

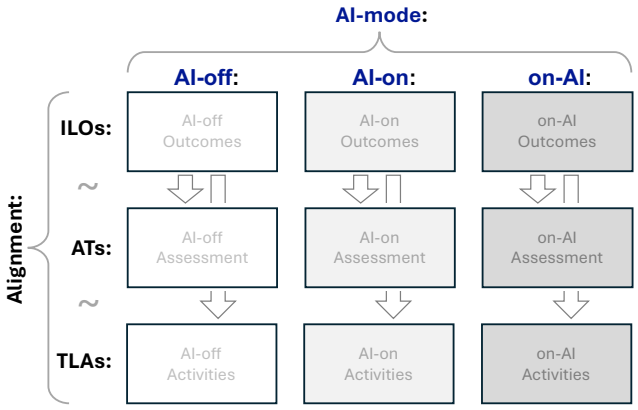

Figure 2: Constructive Alignment extended with an AI-mode dimension. Horizontally, the three side-by-side columns represent the AI-modes (AI-off | AI-on | on-AI), each constituting an instance of Constructive Alignment for that AI-mode. Vertically, alignment occurs within each AI-mode (column), requiring intended learning outcomes (ILOs), assessment tasks (ATs), and teaching/learning activities (TLAs) to be designed consistently, thereby providing aligned learning goals, incentives, and support. Misalignment arises when different AI-modes are mixed across ILOs, ATs, and TLAs.

(1) Formulate Learning & Designate AI-Mode: Formulate the intended learning outcomes (ILOs) and, for each, designate its corresponding AI-mode:

- **Formulate ILOs**
- **Designate AI-mode(ILOs)**

(2) Incentivize Learning: Align Assessment Tasks (ATs) with the Intended Learning Outcomes (ILOs) they are intended to incentivize and ensure that both operate within the same AI-mode:

- **Align ATs ~ ILOs**
- **Align AI-mode(ATs) = AI-mode(ILOs)**

(3) Support Learning: Align Teaching/Learning Activities (TLAs) with the Intended Learning Outcomes (ILOs) they are intended to support and ensure that both operate within the same AI-mode:

- **Align TLAs ~ ILOs**
- **Align AI-mode(TLAs) = AI-mode(ILOs)**

Figure 3: Step-*by*-step process (recipe) for implementing alignment in the age of AI.

and student expectations regarding what the course is intended to achieve.

2.2 Constructive Alignment in the Age of AI

Generative AI does not invalidate this alignment process. However, it makes several long-standing alignment problems considerably more visible; in particular, three challenges emerge:

Outcome Ambiguity (ILOs). What role should AI play in the intended learning? Should students learn *without* AI (AI-off), *with* AI assistance (AI-on), or *about* AI itself (on-AI)?

Assessment Collapse (ATs). Traditional assessments may no longer distinguish autonomous student competence from AI-outsourced competence. Are we assessing the student, an AI system, or some combination of the two?

Activity Drift (TLAs). Teaching/learning activities designed for a *pre*-AI context may no longer prepare students for *post*-AI assessment and professional practice. Conversely, unreflective AI use may reduce opportunities for the deliberate practice required to develop foundational competences.

Problems. The three challenges share a common cause: traditional Constructive Alignment lacks an explicit way of specifying the *intended role of AI* in learning. Without an explicitly designated AI-mode, intended learning outcomes, assessment tasks, and teaching/learning activities may implicitly assume different roles for AI, leading to outcome ambiguity, assessment collapse, and activity drift.

Solution Proposal. To address this limitation, we propose extending Constructive Alignment with an additional design dimension: *AI-mode*. Figure 2 illustrates the resulting framework. Constructive Alignment retains its traditional components—intended learning outcomes (ILOs), assessment tasks (ATs), & teaching/learning activities (TLAs)—but now requires educators to explicitly designate the intended AI-mode of learning. We distinguish between three AI-modes as shown in Table 1.

AI-mode. The central idea is simple. Once an AI-mode has been chosen for an intended learning outcome, the corresponding assessment tasks and teaching/learning activities must be designed consistently within that same AI-mode. Alignment therefore occurs vertically within each column of Figure 2. Note that traditional Constructive Alignment effectively corresponds to the leftmost (white) column, representing learning without AI (AI-off). The middle (light gray) and rightmost (gray) columns instantiate the same principles of Constructive Alignment, but for learning *with* AI assistance (AI-on) and learning *about* AI (on-AI), respectively. Misalignment arises when different AI-modes are mixed across intended learning outcomes, assessment tasks, and teaching/learning activities, corresponding to combinations that span multiple AI-mode columns in Figure 2.

Constructive Alignment in the age of AI can therefore be implemented through the following three-step design process as outlined in Figure 3.

The next three sections examine how this AI-mode dimension reshapes each component of Constructive Alignment, exploring how AI-off, AI-on, and on-AI modes of learning apply to intended learning outcomes (ILOs), assessment tasks (ATs), and teaching/learning activities (TLAs).

3 Intended Learning Outcomes (ILOs)

This section examines the *first* step in the design process:

(1) Formulate Learning & Designate AI-Mode.

Table 1: Summary of AI-modes for Outcomes.

AI-mode:	Learning:	Role of AI:	Guiding question:
AI-off	*without* AI	excluded	Can the student achieve the desired outcome *without* AI?
AI-on	*with* AI	means to an end	Can the student achieve the desired outcome *with* AI?
on-AI	*about* AI	object of learning	Can the student achieve AI literacy: *effective use* of AI?

Outcome Ambiguity is the overarching challenge exposed by AI with respect to *intended learning outcomes*; it arises when the *intended role of AI* in the learning outcome is left *implicit* or *unspecified*. What role should AI play in the intended learning? Are students expected to learn *without* AI assistance, *with* AI assistance, or *about* AI itself (including *when* and *why* not to use it)? When the intended role of AI remains unspecified, educators, students, and assessors may hold different assumptions about what is actually expected. Constructive Alignment in the age of AI therefore requires educators to make the intended role of AI explicit—and to communicate it clearly to students.

In traditional Constructive Alignment, educators formulate Intended Learning Outcomes (ILOs) that specify what students should be able to *do* upon completion of a course. ILOs are therefore typically expressed using observable action verbs; e.g., ***explain***, ***analyze***, ***evaluate***, ***design***, ***implement***, or ***reflect***, paired with an object of learning (verb + noun); e.g., ***analyze*** *algorithms*, ***design*** *experiments*, or ***evaluate*** *evidence*. Taxonomies of cognitive processes, such as SOLO [6] and Bloom [3, 7], provide organizational frameworks and vocabularies that can inspire educators and help them formulate, structure, and communicate intended levels of cognitive complexity.

Intended learning outcomes should specify *observable* competence rather than merely *content* coverage. What matters is not simply what content students encounter, but what they are ultimately expected to *'be able to do'*: *verbs*, not nouns.

AI-mode. In the age of AI, however, specifying the intended competence alone is no longer sufficient. Educators must also specify the *intended role of AI* related to that competence. The challenge identified in Section 2.2 as **Outcome Ambiguity** arises precisely when the role of AI remains unspecified or implicit.

Consequently, we argue that every ILO should be associated with an explicit AI-mode. Just as traditional Constructive Alignment requires educators to make intended learning outcomes explicit, alignment in the age of AI requires educators to *make the intended role of AI explicit*. Without an explicit AI-mode, subsequent alignment of assessment tasks and teaching/learning activities becomes impossible. This transforms questions such as:

"Can the student analyze data?"

into more precise questions, for example:

- **AI-off:** ***analyze*** data *autonomously without* AI assistance;
- **AI-on:** ***analyze*** data *collaboratively with* AI assistance; or
- **on-AI:** *critically* ***review*** AI-generated data analysis output.

The three AI-modes differ according to the intended role of AI in the learning outcome. In AI-off outcomes, AI plays *no* intended role. In AI-on outcomes, AI serves as a legitimate *means* of accomplishing the intended task. In on-AI outcomes, AI itself becomes the *object* of learning. We therefore distinguish between three broad classes of intended learning outcomes:

3.1 AI-off Outcomes

(*aka*, *traditional*, *independent*, or *autonomous* outcomes). **AI-off** outcomes describe competences that students are expected to demonstrate *autonomously*, *without* AI assistance. Such outcomes correspond most closely to traditional pre-AI intended learning outcomes and remain as relevant today as before the emergence of generative AI. In particular, many foundational competences continue to require independent performance, whether for reasons of safety, accountability, transfer of learning, or the development of underlying expertise. Even in domains where AI assistance is expected in professional practice (e.g., computer programming), students may first need to develop sufficient autonomous competence to effectively understand, appreciate, evaluate, adapt, and oversee AI-supported work.

3.2 AI-on Outcomes

(*aka*, *AI-augmented*, *AI-assisted*, or *AI-supported* outcomes). **AI-on** outcomes describe competences that students are expected to demonstrate *with* AI assistance; i.e., intentionally *using* AI. Rather than prohibiting AI, such outcomes explicitly incorporate AI as an intended and legitimate component of competent performance. The focus remains on the quality of the resulting product or performance, but AI is treated as an accepted tool in the work process.

Such outcomes often mirror contemporary professional practice, where increasingly capable AI systems are becoming integrated into everyday work. In many contexts, the educational goal is therefore not merely autonomous performance, but the ability to effectively leverage AI to improve the resulting work. For example, professional software developers increasingly program software with AI assistance, lawyers use AI to summarize and synthesize legal precedents, consultants use AI to develop reports and draft presentations, and marketing professionals use AI to draft and refine communication materials.

Accountability & Responsibility. Importantly, AI-on outcomes do not imply that students relinquish responsibility for their work. Rather, students remain accountable for the quality, correctness, and appropriateness of the resulting product, and the intended competence remains attributable to them, even when AI contributes to the work process. Following the principle of the *Andy's Axe Analogy*, we attribute responsibility to the human rather than the tool—as the analogy goes [2]:

> We say: "*Andy* cut down the tree." We don't say: "An *Axe* cut down the tree." Nor do we say that the tree was cut as a result of an *"Andy–Axe collaboration."* Pressed for more detail, we would say that "Andy

> cut down the tree, *using* the Axe." Andy's *Axe* is just another *tool* [..].

By analogy, the use of increasingly capable AI systems does not transfer ownership or responsibility for the resulting work from the student to the AI system. Students remain responsible both for the appropriate (ab)use of AI and for the resulting consequences of that use. The intended competence therefore lies not merely in operating an AI system, but in effectively *combining* human judgment and AI capabilities to achieve a desired outcome.

Human–AI Collaboration. In educational contexts, students should *add value* to an activity through *meaningful* participation and *competence development*. Although *partial* cognitive outsourcing may be appropriate and even desirable, educational value may diminish when AI use results in *total cognitive outsourcing*, leaving little opportunity for students to exercise or develop the intended competences. In the extreme, students risk becoming *replaceable* participants who contribute little beyond initiating or providing *pro forma* approval of AI-generated outputs—for example, by habitually clicking "OK" or engaging in *Copy-Problem–Copy-Solution* (*CPCS*) workflows without meaningful review or reflection. AI-on learning therefore requires *meaningful* human–AI collaboration and synergy, in which students remain responsible for directing, evaluating, and integrating AI contributions. The educational goal is therefore not merely to keep a *human in the loop*, but to keep them *meaningfully in the loop*:

$$\textit{human in the loop} \implies \textit{human } \boldsymbol{\textit{meaningfully}} \textit{ in the loop.}$$

3.3 on-AI Outcomes

(*aka*, *AI-specific*, *AI-focused*, or *AI-process* outcomes). **on-AI** outcomes describe competences where how-to-use-AI itself becomes the object of learning. Rather than using AI to accomplish a disciplinary task, students are expected to develop knowledge, skills, and judgment concerning AI systems, including their capabilities, limitations, and (in)appropriate (ab)use, and ethical implications. Such outcomes are becoming increasingly important across disciplines as AI systems become embedded in professional and societal practice.

Operational AI Literacy. Recent work has argued that the age of AI shifts emphasis from artifact production towards higher-level competences concerned with specifying, evaluating, and supervising AI-generated work [1, 8, 10]. Building on this work, researchers in computing education have identified three broad classes of *operational AI-literacy competences*:

- **Prompt:** Formulate clear and unambiguous input specifications that guide AI towards generating intended artifacts;
- **Review:** Critically assess AI-generated artifacts to ensure alignment with intended purpose, without unintended consequences;
- **Orchestrate:** Supervise AI and manage the interplay between human activities and AI-generated contributions.

The prompt, orchestrate, and review competence triad provides an operationalization of AI literacy that is teachable, assessable, and actionable in educational practice. Further, these competences are largely generic and transferable across domains, although their manifestation depends on disciplinary norms and epistemic practices [1, 8, 10]. For example, reviewing AI-generated software requires different forms of expertise than reviewing AI-generated legal analyses, educational materials, or marketing content. The operational AI-literacy competences discussed here represent only one class of on-AI competences. Courses may also address:

- **Conceptual understanding of AI** (e.g., capabilities, limitations, and failure modes); *and*
- **Ethical reasoning about AI** (e.g., bias, fairness, privacy, and accountability).

Discipline-specific AI competences. In addition to generic AI competences, disciplines may define discipline-specific on-AI outcomes. For example, a statistics course might include the outcome ***evaluate*** AI-generated data analysis output, while a programming course might require students to ***review*** AI-generated code for correctness, quality, security, maintainability, inclusion, and absence of bias.

3.4 Example Outcomes

To illustrate the proposed framework, Table 2 presents example intended learning outcomes (ILOs) for a *Research Methods* course. This course was chosen intentionally because it represents a broad, area-independent subject that is commonly found across disciplines, thereby illustrating the general applicability of the framework beyond any particular field of study. The purpose of the example is not to prescribe a particular set of outcomes, but to showcase how AI-mode can be associated with individual outcomes (ILOs).

The example deliberately includes outcomes from all three AI-modes. The first two outcomes focus on competences that students are expected to demonstrate autonomously without AI assistance (AI-off), such as the ability to *explain* fundamental research methods or *design* an empirical study. The next two outcomes explicitly incorporate AI as an intended and legitimate component of competent performance (AI-on), such as the ability to *develop* research protocols or *analyze & visualize* data with AI assistance. Finally, the last two outcomes focus on AI itself as the object of learning (on-AI), such as *evaluating* and *refining prompting* strategies for generating research study artifacts or *identifying* and *assessing* hallucinated or invalid claims in AI-generated *reviews* of research designs.

Importantly, AI-mode is a property of individual intended learning outcomes rather than of an entire course. Consequently, courses may consist exclusively of AI-off, AI-on, or on-AI outcomes, or they may combine multiple AI-modes within the same course. The *Research Methods* course in Table 2 illustrates the latter case, incorporating outcomes from all three AI-modes. The framework does not require any particular combination of AI-modes. It merely requires that each intended learning outcome be assigned an explicit AI-mode, and that its corresponding assessment tasks and teaching/learning activities be aligned consistently with that mode.

4 Assessment Tasks (ATs)

This section examines the *second* step in the design process:

(2) Incentivize Learning.

Assessment Collapse is the overarching challenge exposed by AI with respect to *assessment tasks*; it arises when assessment tasks

Table 2: Example intended learning outcomes (ILOs) for a Research Methods course, tagged by AI-mode.

After the course, students should be able to...:

Intended Learning Outcome (ILO):	**AI-mode:**
• ***Explain*** fundamental research methods (without the use of AI).	AI-off
• ***Design*** an empirical study to address a research question (without the use of AI).	AI-off
• ***Develop*** research protocols using AI assistance.	AI-on
• ***Analyze*** and ***Visualize*** data using AI assistance.	AI-on
• ***Evaluate*** and ***Refine*** *prompting* strategies for generating research study artifacts.	on-AI (*Prompt*)
• ***Identify*** and ***Assess*** hallucinated or invalid claims in AI-generated *reviews* of research designs.	on-AI (*Review*)

no longer provide valid evidence of the intended competence. Traditional assessment tasks were originally designed under the assumption that students would perform work autonomously. With the emergence of generative AI, however, that assumption can no longer be taken for granted. Consequently, assessments intended to measure one competence may instead measure a different competence, or fail to distinguish between them altogether.

This challenge is particularly important because assessment plays a dual role within Constructive Alignment. Assessment not only *measures* learning; it also *incentivizes* learning. Although assessment is often encountered at the *end* of a course from an educator's perspective, it effectively appears at the *beginning* from the student's perspective because it shapes how students choose to approach their learning [5]. Students naturally direct their efforts towards what is assessed. When assessment becomes misaligned with the intended learning outcomes, students may therefore optimize for success on the assessment rather than for developing the intended competence.

Assessment Collapse is *effectively* a problem of **AI-mode misalignment** between intended learning outcomes (ILOs) and assessment tasks (ATs): ILOs $\not\sim$ ATs. The most problematic form of misalignment is illustrated below as a simplified 3×3 square-cluster visualization of the structure introduced in Figure 2:

(Assessment Collapse) 

Consistent with Figure 2, the *three rows* correspond to ILOs, ATs, and TLAs, while the *three columns* correspond to AI-off, AI-on, and on-AI (the three AI-modes). The form of AI-mode misalignment most commonly associated with *Assessment Collapse* is highlighted by the red square (■): an AI-on assessment task misaligned with an AI-off intended learning outcome; i.e.:

(■) AI-**off** ILO $\not\sim$ AI-**on** AT (■).

Here, the intended learning outcome is that students *develop* autonomous competence (AI-off), while the assessment task *permits*—perhaps even *rewards*—AI-assisted performance (AI-on). This is precisely the worrying scenario that currently concerns many institutions and educators. Students may succeed on the assessment without demonstrating the competence that the assessment was intended to measure, thereby undermining the *validity* of the assessment as evidence of that competence. In such situations, what is effectively assessed can collapse toward AI-mediated production rather than independent reasoning. The result is an *indirect* failure to acquire the intended competence, driven not by lack of ability, but by lack of *incentive* to engage in the targeted practice.

Of course, this phenomenon predates generative AI. Long before AI, students could seek assistance from friends, family members, tutors, online resources, previous reports, and numerous other external sources of support. In principle, such assistance could similarly undermine the relationship between the competence intended to be assessed and the competence actually demonstrated. What distinguishes generative AI is not the existence of external assistance, but its unprecedented accessibility, immediacy, scalability, and perceived legitimacy. As a result, forms of assessment misalignment that were previously limited, costly, or difficult to detect have become both widespread and impossible to ignore.

Importantly, AI does not create this problem. Rather, AI exposes and intensifies a pre-existing assessment misalignment by making cognitive outsourcing inexpensive, immediate, and widely accessible. AI does not cause cheating; it reveals assessments that were already vulnerable to misalignment.

The appropriate response in this worrying scenario is to assess autonomous intended learning outcomes through autonomous (AI-off) assessment, as indicated by the green square (■):

(Assessment Collapse)  ⇒  **(Aligned Assessment)**

thereby ensuring AI-mode alignment

(■) AI-**off** ILO $\sim$ AI-**off** AT (■).

More generally, the solution is not necessarily to prohibit AI. Nor is it necessarily to require AI. Rather, the solution is to simply align the AI-mode of assessment with the AI-mode of the intended learning outcome.

Other forms of AI-mode misalignment are possible; in particular:

(Misaligned ILO $\not\sim$ AT) 

This 3×3 square cluster illustrates a different form of AI-mode misalignment in which students are intended to develop AI-assisted competence (AI-on), but are assessed autonomously, without access to AI (AI-off), represented by the dark red square (■) In this case, the assessment may fail to measure the competence that the course intended students to develop. Students who have successfully learned to work effectively with AI may therefore be disadvantaged by an assessment that instead assumes and rewards autonomous performance.

Once again—assuming the intended learning outcomes correctly capture the intended competences—the CA solution is to *align* the AI-mode of assessment with the AI-mode of the outcomes, as indicated by the dark green square (■):

(Misaligned ILO ≁ AT)  ⇒  **(Aligned ILO ~ AT)**

The first misalignment example illustrated *Assessment Collapse*, where assessment no longer provides valid evidence of the intended competence and may consequently lead to an *overestimation* of competence. In contrast, the second misalignment example may lead to an *underestimation* of competence by assessing students under conditions inconsistent with the intended learning.

More generally, AI-mode misalignments can create confusion about course expectations and send conflicting signals about what is expected, valued, and rewarded. Such misalignments may invalidate assessments as evidence of the intended competence and cause students to direct their efforts toward competences other than those intended by the course.

We therefore distinguish between three broad classes of assessment tasks:

4.1 AI-off Assessment

(aka, *autonomous*, *invigilated*, *proctored*, or *protected* assessment). **AI-off** assessment tasks evaluate competences demonstrated *without* AI assistance. Such assessments are appropriate when the intended learning outcomes themselves target autonomous competence and students are therefore expected to perform independently.

Within Constructive Alignment, the purpose of AI-off assessment is not to prohibit AI per se, but to assess and obtain *valid evidence* of the intended competence. When the intended learning outcome is AI-off, the assessment task must likewise be AI-off.

Examples of AI-off assessments include invigilated or proctored written examinations without access to AI tools, oral examinations, live demonstrations of competence, and justification-focused tasks requiring students to explain and defend their reasoning, all without AI-mediated assistance. Such assessments often provide particularly strong evidence of competence because students must articulate their understanding, make decisions, and respond to follow-up questions in real time, without AI access.

Oral examinations may be especially robust in the age of AI, as long as students cannot receive AI-mediated assistance during the assessment. By requiring students to demonstrate competence interactively and without external assistance, oral assessments align incentives for autonomous practice during learning. More generally, assessment *incentivizes* students to prepare under the AI-unassisted conditions in which they expect to be assessed.

Importantly, AI-off assessments (and, by alignment, AI-off outcomes) do not imply that AI should be absent from teaching/learning activities. Rather, they indicate that the competence being assessed is intended to be demonstrated autonomously (without AI).

4.2 AI-on Assessment

(often implemented through *authentic assessment*). **AI-on** assessment tasks evaluate competences demonstrated *with* AI assistance. Such assessments are appropriate when the intended learning outcomes themselves target (intended) AI-assisted competence and students are therefore expected to work *with* AI tools.

Within Constructive Alignment, the purpose of AI-on assessment is not merely to permit AI use, but to obtain *valid evidence* of students' ability to *use* AI appropriately and effectively in accomplishing disciplinary tasks. When the intended learning outcome is AI-on, the assessment task must likewise be AI-on.

Examples of AI-on assessments include take-home projects, portfolio work, design tasks accompanied by reflection, and other assessments in which AI use is explicitly *permitted* or even *required*. Such assessments often focus on the quality of the resulting *product* or performance rather than the *process* of AI use itself, and frequently resemble professional practice, where competent work increasingly involves effective collaboration with AI tools.

In AI-on assessment, AI functions as a *means* to an end rather than an *end* in itself. Consistent with *Andy's Axe Analogy* (Section 3.2), tools do not receive responsibility or credit for the resulting work. Responsibility therefore ultimately remains with the student rather than the AI system. The competence being assessed is not merely the operation of AI tools, but the ability to achieve disciplinary goals through the effective integration of human judgment and AI capabilities.

Human–AI Collaboration. Because AI-on intended learning outcomes emphasize *meaningful* human–AI collaboration, assessments should elicit evidence of *meaningful human involvement* rather than superficial AI use (e.g., *Copy-Problem–Copy-Solution*). Assessments that permit AI-generated outputs without requiring students to exercise judgment, oversight, or responsibility may provide weak evidence of the intended competence. Students should therefore remain *meaningfully in the loop*, demonstrating how they direct, evaluate, integrate, and govern AI contributions. In this way, assessments remain aligned with intended learning outcomes centered on *meaningful* human–AI collaboration.

4.3 on-AI Assessment

(aka, *AI-process*, *AI-literacy*, or *AI-use* assessment). **on-AI** assessment tasks evaluate students' understanding and use of AI itself. Such assessments are appropriate when AI forms part of the *object* of learning itself rather than merely a means of accomplishing disciplinary tasks.

In contrast to AI-on assessment, which often emphasizes the quality of the resulting *product* or *performance*, on-AI assessment primarily focuses on the *process* of AI use itself—in short, *process over product*.

Examples include assessments of student ability to *prompt*, *orchestrate*, and *review* AI systems and their outputs. Assessment may focus on formulating effective prompts, supervising human–AI workflows, critically evaluating AI-generated artifacts, and selecting appropriate AI system(s) for particular task(s). Students may further be asked to identify hallucinations or biases, compare alternative AI outputs, document and justify decisions regarding the (non-)use of AI, critically evaluate the capabilities and limitations of AI systems, and reason about ethical issues such as fairness, accountability, and privacy.

Table 3: Summary of AI-modes for Assessment.

AI-mode:	Assessment focus:	Typical formats:	Guiding question:
AI-off	autonomous competence	oral exams, invigilated exams	Can the student do it *without* AI?
AI-on	product/performance	projects, portfolios	Can the student do it *with* AI?
on-AI	AI process/judgment	prompting, think-aloud AI use	Can the student (non-)*use* AI *effectively* and *responsibly*?

Suitable assessment formats include live interactive demonstrations of AI use, think-aloud sessions, oral examinations, and submission of authentic interaction histories with AI systems. Dynamic assessments may provide particularly strong evidence of competence, as static artifacts such as prompts or chat logs can themselves be generated or manipulated using AI. Some AI systems additionally allow original interaction histories to be shared or verified, thereby providing stronger evidence of authentic AI use.

The competence being assessed is therefore not the disciplinary *product* itself, but rather student ability to understand, evaluate, and employ AI appropriately and critically throughout the *process* of (non-)working with AI.

4.4 Summary of AI-modes for Assessment

Table 3 summarizes the three AI-modes as they apply to assessment tasks. Consistent with the distinction introduced for intended learning outcomes, the three AI-modes differ not only in whether AI is used, but also in the primary focus of assessment. Whereas AI-off assessment emphasizes autonomous competence demonstrated *without* AI, AI-on assessment often focuses on the quality of the resulting product or performance achieved *with* AI. In contrast, on-AI assessment primarily focuses on students' processes and judgments regarding the (non-)use of AI itself.

4.5 What about an AI-assessor?

(aka, *automated*, *AI-assisted*, or *AI-based* assessment). The preceding discussion has focused on AI as part of students' learning and assessment. A natural question is whether AI might also act as an *assessor* (or, *evaluator*) itself. Although still emerging in educational contexts, AI systems are increasingly being used to support or automate evaluative decisions in domains such as recruitment, candidate screening, and performance assessment [31, 34].

At the *formative* end of the spectrum, students may themselves use AI systems to simulate examinations, provide rubric-based feedback, or estimate likely grades on assignments. For example, students preparing (training) for an oral examination might interact with an AI system that poses questions, evaluates responses, and provides feedback on their current level of competence. Such forms of AI-assisted self-assessment may support learning by helping students identify weaknesses and prepare for subsequent assessment.

At the *summative* end of the spectrum, AI systems could conceivably assist educators in grading assignments, screening large numbers of submissions, or even conducting interactive assessments. Students might interact with an AI-assessor that probes understanding and ultimately produces an evaluative judgment or even a final grade. Similarly, organizations may increasingly employ AI systems to screen applicants before more detailed human evaluation.

Such uses, however, raise important ethical questions regarding validity, reliability, transparency, bias, accountability, and opportunities for appeal. Constructive Alignment remains relevant in this context: regardless of whether assessment is performed by humans, AI systems, or some combination thereof, assessments must still provide *valid evidence* of the intended learning outcomes.

The question is therefore not whether AI can act as an assessor, but whether AI-assisted assessment can validly and fairly measure the competences that educators intend students to develop.

5 Teaching/Learning Activities (TLAs)

This section examines the *third* and final step in the design process:

(3) Support Learning.

Activity Drift is the overarching challenge exposed by AI with respect to *teaching/learning activities*; it arises when teaching/learning activities no longer *support* the intended learning process. Before the age of AI, teaching/learning activities were often designed under the assumption that students would engage with them as intended. With the emergence of generative AI, that assumption can no longer be taken for granted. Students may now *bypass* learning activities entirely through "cognitive outsourcing" to AI systems, or engage with activities in other ways that are inconsistent with the intended learning outcomes.

This challenge is particularly important because teaching/learning activities provide the primary means through which students develop the intended competences. They do so by providing opportunities for guided practice, formative feedback, and competence development. When activities become misaligned with the intended learning outcomes, students may therefore fail to develop the competences that the course intends them to acquire.

Activity Drift is *effectively* a problem of **AI-mode misalignment** between intended learning outcomes (ILOs) and teaching/learning activities (TLAs): ILOs $\not\sim$ TLAs. The most problematic form of misalignment is illustrated below as a simplified 3×3 square-cluster visualization of the structure introduced in Figure 2:

(Activity Drift) 

Consistent with Figure 2, the *three rows* correspond to ILOs, ATs, and TLAs, while the *three columns* correspond to AI-off, AI-on, and on-AI (the three AI-modes). The most problematic form of AI-mode misalignment is illustrated below. The red square (■) highlights an AI-on teaching/learning activity that is misaligned with an AI-off intended learning outcome:

(■) AI-**off** ILO $\not\sim$ AI-**on** TLA (■).

Here, the intention is that students develop autonomous competence, but teaching/learning activities that are insufficiently aligned

with this intention may permit—or even implicitly encourage—students to outsource the activity to AI systems without engaging in the intended learning process, for example through a *Copy-Problem–Copy-Solution* (CPCS) workflow. Consequently, students may complete the activity without engaging in the deliberate practice necessary to develop the intended autonomous competence. When both *assessment* and *activities* (Sections 4 & 5) are appropriately aligned with the intended learning outcomes, and *activities* are realistically doable (designed within students' zones of proximal development (ZPD) [33]), students have stronger *incentives* and *support* to properly engage in the intended learning process rather than resorting to AI-mediated completion.

Whereas assessment misalignment primarily threatens the *validity of evidence* of competence, activity misalignment threatens the *acquisition of competence* itself. The result is therefore a *direct* failure to acquire the intended competences through practice.

Of course, this phenomenon (lack of learning engagement) predates generative AI. Long before AI, students could rely on friends, tutors, solution manuals, online resources, or previous reports rather than engaging directly with learning activities. As mentioned earlier, what distinguishes generative AI is not the existence of external support, but its unprecedented accessibility, immediacy, scalability, and perceived legitimacy. As a result, forms of activity misalignment that were previously limited, costly, or difficult to detect have become both widespread and impossible to ignore.

Importantly, AI does not create this problem. Rather, AI exposes and intensifies a pre-existing misalignment between intended learning outcomes and the teaching/learning activities intended to support them. AI does not *cause* poor learning design; it *reveals* learning activities that never effectively supported the intended competence in the first place.

Whenever assessment is subsequently aligned with the intended learning outcomes, students who have not acquired the intended competences through practice risk failing the assessment down the road. When assessment is itself misaligned in the same way, however, students may instead pass the course without ever acquiring the intended competences.

The solution is not necessarily to *prohibit* AI. Nor is it necessarily to *require* AI. Rather, the solution is to *align* the AI-mode of teaching/learning activities with the AI-mode of the intended learning outcome, as indicated by the green square (■):

(Activity Drift)  ⇒  **(Aligned Activity)**

thereby ensuring AI-mode alignment:

(■) AI-**off** ILO ~ AI-**off** TLA (■)

Other forms of AI-mode misalignment are possible; in particular:

(Misaligned ILO ≁ TLA) 

This 3×3 square cluster illustrates a different form of AI-mode misalignment in which students are intended to develop *AI-assisted* competence (AI-on), but are provided only with activities that *prohibit* or *exclude* AI use (AI-off). In this case, students may receive insufficient opportunities to *develop* the AI-assisted competences that the course intends them to acquire.

Once again—assuming the intended learning outcomes correctly capture the intended competences—the CA solution is to *align* the AI-mode of teaching/learning activities with the AI-mode of the outcomes, as indicated by the dark green square (■):

(Misaligned ILO ≁ TLA)  ⇒  **(Aligned ILO ~ TLA)**

Whereas the previous example may lead to *under-training* through insufficient autonomous practice, this form of misalignment may instead lead to *under-preparation* for professional contexts in which effective AI use is expected.

More generally, AI-mode misalignments can create confusion about course expectations and deprive students of appropriate opportunities to develop the intended competences. Such misalignments may therefore undermine learning even when assessment remains perfectly aligned.

We therefore distinguish between three broad classes of teaching/learning activities:

5.1 AI-off Activities

(aka, *autonomous*, *independent*, or *deliberate-practice* activities). **AI-off** teaching/learning activities *support* competences developed *without* AI assistance. Such activities are appropriate when the intended learning outcomes themselves target autonomous competence and students are therefore expected to learn and practice independently.

Within Constructive Alignment, the purpose of AI-off activities is not to *prohibit* AI per se, but to *support* the development of autonomous competence. When the intended learning outcome is AI-off, the teaching/learning activity must likewise be AI-off.

The primary purpose of AI-off activities is not merely task completion, but competence development through deliberate practice. Exposure to content does not necessarily imply competence development. Learning *about* a practice is not equivalent to learning *to perform* that practice. After all, *learning about cooking* is not the same as *learning to cook*. Such activities provide opportunities for students to engage in the cognitive processes through which competence is acquired. Importantly, AI-off activities should ideally be designed within students' zones of proximal development (ZPD), providing sufficient challenge to promote learning while reducing the likelihood of students resorting to AI-mediated completion of the activities.

Competence Development. This concern is particularly important and apparent in high-stakes domains. Few would feel comfortable being operated on by a surgeon who had never autonomously developed surgical competence, or boarding a plane flown by a pilot who had never autonomously practised aircraft emergency control procedures. The concern is not tool use itself, but whether students have sufficient opportunities to *develop* the autonomous competences required for safe and accountable professional practice. AI is not the problem; lack of autonomous competence development is. The concern here is systemic rather than individual: the risk lies not in tool use itself, but in a misaligned educational system that inadvertently fails to provide the practice required for safe and effective professional practice.

Examples of AI-off activities include *exercises*, *problem-solving sessions*, *live demonstrations*, *whiteboard reasoning*, *retrieval practice*, *peer instruction*, and other activities in which students actively

engage with disciplinary content without AI assistance. Such activities intentionally preserve *productive cognitive friction*, requiring students to struggle, reflect, and reason through problems as part of the learning process.

Importantly, AI-off activities do not imply that AI should be absent from the course as a whole; they merely indicate that this *particular* competence being developed is intended to be acquired autonomously, *without* AI assistance. Other intended learning outcomes within the same course may legitimately employ AI-on or on-AI activities.

In AI-off activities, the educational goal is not to collapse the distance between problem and solution, but to help students traverse that distance themselves.

5.2 AI-on Activities

(aka, *AI-assisted*, *AI-augmented*, or *AI-mediated* activities). **AI-on** teaching/learning activities support competences developed through collaboration with AI systems. Such activities are appropriate when the intended learning outcomes themselves target AI-assisted competence and students are therefore expected to learn and practice *with* AI.

Within Constructive Alignment, the purpose of AI-on activities is not merely to expose students to AI technologies, but to support the development of meaningful and effective human–AI collaboration. When the intended learning outcome is AI-on, the teaching/learning activity must likewise be AI-on to *foster* such competence development.

The primary purpose of AI-on activities is not autonomous competence development, but competence development through effective human–AI collaboration. Such activities often resemble professional practice, where competent performance increasingly involves effective collaboration with AI tools.

Examples of AI-on activities include *prompt refinement workshops*, *comparative solution analyses* (e.g., *human versus AI solutions*), *AI-assisted design activities*, *iterative prompting exercises*, and *critical review of AI-generated artifacts*. Such activities provide opportunities for students to develop professional competences for working effectively with AI systems while maintaining responsibility for the resulting work.

Importantly, in AI-on activities, AI functions as a *means* to an end rather than an *end* in itself. The educational goal is therefore not merely to use AI, but to use AI effectively and responsibly in pursuit of disciplinary goals.

A useful distinction between AI-on and on-AI activities is that the former primarily focuses on the resulting *product* or performance, whereas the latter primarily focuses on the *process* of AI *use* itself. In AI-on activities, AI is treated as a means toward achieving disciplinary goals; in on-AI activities, the use of AI itself becomes the object of learning.

As with AI-off activities, AI-on activities need not characterize an entire course. Rather, they are associated with *particular* intended learning outcomes, while other outcomes within the same course may legitimately employ AI-off or on-AI activities.

Human–AI Collaboration. Because AI-on intended learning outcomes emphasize *meaningful* human–AI collaboration, teaching/learning activities should support such collaboration and synergy rather than unreflective delegation to AI systems. Activities that reduce students to passive acceptance of AI-generated outputs may provide little support for competence development. Teaching/learning activities should therefore be designed to keep students *meaningfully in the loop* throughout the learning process. In this way, teaching/learning activities remain aligned with intended learning outcomes centered on *meaningful* human–AI collaboration.

5.3 on-AI Activities

(aka, *AI-process*, *AI-literacy*, or *AI-specific* activities). **on-AI** teaching/learning activities support competences in which AI itself constitutes the object of learning. Such activities are appropriate when the intended learning outcomes target students' ability to understand, choose, evaluate, govern, and effectively use AI systems.

Within Constructive Alignment, the purpose of on-AI activities is not merely to use AI technologies, but to develop competences *about* AI. When the intended learning outcome is on-AI, the teaching/learning activity must likewise be on-AI to foster such competence development.

The primary purpose of on-AI activities is not the production of disciplinary artifacts, but the development of competences related to the *use* of AI itself. In particular, such activities provide opportunities for students to develop operational AI literacy competences through learning how to effectively ***prompt***, ***orchestrate***, and ***review*** AI systems and their outputs (cf. Section 3.3).

Examples of on-AI activities include *prompt engineering exercises*, *think-aloud AI interactions*, *workflow orchestration tasks*, *hallucination detection*, *bias identification*, *model comparison*, *critical review of AI-generated artifacts*, *ethical analysis of AI use*, and *reflection on the appropriate (non-)use of AI*. Students may further be asked to *document* and *justify decisions* regarding the use of AI systems, including which system(s) are appropriate for particular task(s).

Process over Product. Importantly, whereas AI-on activities primarily focus on the resulting *product* or performance, on-AI activities primarily focus on the *process* of AI *use* itself. In on-AI activities, AI is not merely a means toward achieving disciplinary goals, but itself becomes the object of learning.

Because the focus is on the process of AI use, dynamic activities may provide particularly strong evidence of competence. Static artifacts such as *prompts* or *chat logs* may themselves be generated or manipulated using AI, although some AI systems allow authenticated sharing of original interaction histories.

As with AI-off and AI-on activities, on-AI activities need not characterize an entire course. Rather, they are associated with *particular* intended learning outcomes, while other outcomes within the same course may legitimately employ AI-off or AI-on activities.

5.4 Summary of AI-modes for Activities

Table 4 summarizes the three AI-modes for teaching/learning activities. As with intended learning outcomes and assessment tasks, the AI-mode of each teaching/learning activity should ultimately

Table 4: Summary of AI-modes for Activities.

AI-mode:	Role of AI:	Activity focus:	Typical activities:	Guiding question:
AI-off	excluded	autonomous practice	exercises, live demonstrations	How can students learn *without* AI?
AI-on	means to an end	product/performance	prompt refinement, AI-assisted design	How can students learn *with* AI?
on-AI	object of learning	AI (non-)use process	prompt, orchestrate, review	How can students learn *about* AI (non-)*use*?

be determined by the competence students are expected to acquire. The goal is not to prescribe a single AI-mode for an entire course, but to align the AI-mode of each activity with the intended learning outcome it is designed to support.

5.5 What about an AI-tutor?

Although inherently based on AI, an AI-tutor differs fundamentally from an AI-assistant designed primarily for task completion. The primary purpose of an AI-tutor is not to generate solutions, but to support the *learning process* through explanation, questioning, feedback, and scaffolding. In this sense, an AI-tutor resembles a personal *human* tutor more than an answer-generating system. This distinction is captured by the familiar proverb:

> *"Give a person a fish and you feed them for a day; teach a person to fish and you feed them for a lifetime."*
>
> — Chinese proverb

Problem–Solution Distance. The educational value of an AI-tutor lies not in collapsing the distance between *problem* and *solution*, but in facilitating (helping) students traverse that distance themselves. Accordingly, it should primarily operate in the space between *problem* and *solution*, supporting learning through explanation, questioning, hints, feedback, retrieval support, and scaffolding rather than directly providing solutions.

An AI-tutor may support learning in numerous ways, including:

- **Feedback & Diagnosis:**
  - *Provide* personalized feedback;
  - *Monitor* progress;
  - *Identify* weaknesses.
- **Instructional Support:**
  - *Retrieve* content;
  - *Explain* concepts;
  - *Facilitate* student understanding of problems/tasks.
- **Practice & Reinforcement:**
  - *Generate* additional exercises for deliberate practice;
  - *Consolidate* newly acquired knowledge;
  - *Generalize* knowledge to new contexts;
  - *Enable* low-stakes practice and experimentation.
- **Motivation & Scaffolding:**
  - *Manage* expectations;
  - *Motivate*, *encourage*, and *support* learners emotionally;
  - *Reduce* scaffolding gradually and transfer responsibility to the learner as competence develops.

An important advantage of an AI-tutor is that students may ask the same "dumb questions" repeatedly without fear of embarrassment, social judgment, or exhausting the patience of the tutor.

Orthogonal to AI-mode. Because an AI-tutor primarily supports the *learning process* rather than disciplinary task completion, it is largely orthogonal to *AI-mode* and may therefore support learning across all three *AI-modes*. For example, an AI-tutor may help students understand concepts and practice autonomously in AI-off contexts, support effective human–AI collaboration in AI-on contexts, and provide opportunities to practice prompting, orchestration, and review in on-AI contexts.

More generally, an AI-tutor may support students throughout the entire Constructive Alignment process. It may help students interpret intended learning outcomes (ILOs), prepare for assessment tasks (ATs), and engage with teaching/learning activities (TLAs). In this sense, an AI-tutor complements AI-assessors (Section 4.5): whereas an AI-assessor evaluates learning, an AI-tutor supports it.

Recent courses have begun integrating AI-tutors directly into the learning environment. For example, Harvard's CS50 course [17] has introduced an AI-tutor (called the "CS50 Duck"), extending the course's long-standing pedagogical tradition of encouraging students to reason through problems rather than simply providing solutions.

Like all educational technologies, however, the value of an AI-tutor ultimately depends not on the technology itself, but on how it is aligned with the intended learning outcomes.

6 Relationship to Existing Work

Recent work on AI in education has largely focused on AI literacy, assessment reform, or AI-enabled teaching tools. In contrast, relatively little attention has been given to extending foundational educational theory to account explicitly for AI.

The proposed AI-mode framework builds on these complementary strands of work while addressing a different question: how AI can be incorporated into the theory of Constructive Alignment itself. The following subsections relate the framework to existing work on Constructive Alignment, AI literacy, assessment reform, and AI tutors and assessors, highlighting both points of convergence and the distinct contributions of the present work.

6.1 Constructive Alignment

Constructive Alignment (CA) provides the theoretical foundation for our argument. Biggs introduced Constructive Alignment as a way of connecting constructivist learning theory with instructional design [4], and it has since become a central principle in outcome-based curriculum design [15]. Its core contribution is to shift attention from what teachers cover to what students are expected to do, how that learning is assessed, and how teaching/learning activities support the intended learning.

At the same time, prior work cautions that CA can be weakened when reduced to a compliance exercise. Kandlbinder emphasizes that alignment should not be treated merely as arranging curriculum components into a consistent pattern [15], while Loughlin,

Lygo-Baker, and Lindberg-Sand argue that CA has often been repurposed as a quality assurance mechanism rather than used as a professional tool for improving teaching and learning [21]. Empirical work further suggests that students' perceptions of alignment matter, and that clarity of intended learning outcomes, alignment of teaching/learning activities and assessment tasks, and effective feedback are associated with motivation, perceived usefulness, and deeper learning strategies [29].

Generative AI makes these alignment questions more urgent because it can decouple task completion from the cognitive work that learning activities and assessments are intended to elicit. Santschi makes a closely related argument, proposing that AI policy should be anchored in tasks rather than tools, and that decisions about appropriate AI use depend on the cognitive work an activity requires and how AI changes that work [30]. We take this one step further by making the role of AI part of alignment itself.

6.2 AI Literacy and AI Competence Frameworks

In our extension of Constructive Alignment, the *on-AI* mode concerns intended learning outcomes in which AI itself is the object of learning. This mode is closely connected to recent work on AI literacy and AI competence, which argues that students need explicit competencies for understanding, using, and critically evaluating AI systems. Ng et al. conceptualize AI literacy as involving the ability to *know & understand* AI, *use & apply* AI, *evaluate & create* with AI, and *address* ethical issues associated with AI systems [23]. Similarly, UNESCO's student AI competency framework defines AI competence across four dimensions: *human-centred mindset*, *ethics of AI*, *AI techniques & applications*, and *AI system design* [22]. The OECD and European Union likewise frame AI literacy as a combination of *knowledge*, *skills*, and *attitudes* that enables learners to *engage with*, *create with*, *manage*, and *shape* AI while critically evaluating its *benefits*, *risks*, and *ethical implications* [24].

These frameworks support the view that AI use in education cannot be reduced to access to tools. Students need to learn how to use AI *critically*, *responsibly*, and *effectively*. The prompt–review–orchestrate competencies that are emerging in computing education research provide one operationalization of this idea as students must learn to *formulate* effective inputs [27], critically *evaluate* AI-generated outputs [28], and *manage* human–AI workflows. In our framework, these are examples of *operational on-AI* outcomes because the intended competence is not simply completing a disciplinary task with AI, but developing judgement about AI (non-)use itself.

AI literacy must also include judgement about when AI use is educationally (in)appropriate. Lodge and Loble distinguish *beneficial cognitive offloading* from *detrimental outsourcing*, arguing that AI can support learning when it *augments cognition* but may undermine learning when it allows students to *bypass the cognitive effort needed for durable understanding* [20]. This distinction reinforces the need for AI-mode alignment. Some outcomes require students to learn *about* AI, some require them to work *with* AI, and others still require practice *without* AI so that foundational competence can develop.

6.3 Assessment Reform in the Age of AI

A substantial body of recent work has focused on *assessment reform* in response to generative AI. The TEQSA report on assessment reform argues that generative AI has exacerbated existing assessment challenges rather than creating them from scratch, especially because non-invigilated assessment is difficult to secure and AI use is difficult to detect reliably [19]. Rather than relying on blanket bans or attempting to secure every assessment task, the report calls for authentic engagement with AI, program-level assessment design, evidence of learning processes, appropriate collaboration with people and AI, and targeted security at meaningful points across a program [19]. This aligns with our argument that AI acts as a *catalyst*: it exposes and intensifies existing weaknesses in assessment design and therefore requires a systemic educational response.

Other work has proposed practical tools for clarifying permitted AI use in assessment. Perkins et al. introduce the Artificial Intelligence Assessment Scale (AIAS), a five-level framework for communicating expectations regarding AI use in assessment. The current version distinguishes five categories[2] [25, 26]:

(1) *No AI*,
(2) *AI Planning*,
(3) *AI Collaboration*,
(4) *Full AI*, and
(5) *AI Exploration*.

The AIAS is valuable because it moves beyond a simple *allowed*-versus-*prohibited* binary and provides educators and students with clearer language and taxonomy for specifying acceptable AI use in assessment. It also foregrounds *transparency*, *academic integrity*, and *student responsibility*.

Our framework is complementary to, but distinct from, these assessment-focused approaches. TEQSA provides sector-level principles for assessment reform, while AIAS provides a practical taxonomy for communicating permitted AI use in assessment tasks. By contrast, AI-mode is an alignment dimension spanning *intended learning outcomes*, *assessment tasks*, and *teaching/learning activities*. The two approaches are therefore *orthogonal* and can be *combined* naturally: AI-mode determines the appropriate role of AI within an assessment, whereas AIAS specifies the permitted forms of AI use within that role.

A related response is the *Two-Lane Approach* to assessment [11]. In this model, Lane 1 assessments are primarily assessments *of* learning: they are typically secured, conducted in person, and used to validate students' attainment of relevant learning outcomes at selected points in a program. Lane 2 assessments are primarily assessments *for* and *as* learning: they are typically more open, often operate at the unit or course level, and may scaffold students' productive and responsible use of AI. This distinction overlaps with the AI-off and AI-on modes proposed in this paper, but it is organized around assessment purpose, security, and program-level assurance rather than around the intended role of AI in the competence itself.

Our framework therefore extends these approaches in three ways. First, we treat AI-mode as a property of individual intended

[2]The original AIAS (Version 1) defined five different levels (*No AI Use*, *AI-Assisted Idea Generation*, *AI-Assisted Editing*, *AI Task Completion with Human Evaluation*, and *Full AI Use*), but these were revised in AIAS Version 2 based on implementation experience.

Table 5: Summary of AI-modes for Constructive Alignment. Traditional alignment corresponds to the (top) AI-off row.

AI-mode:	Learning:	Role of AI:	ILOs:	ATs:	TLAs:
AI-off	*without* AI	excluded	autonomous competence	autonomous performance	autonomous practice
AI-on	*with* AI	means to an end	AI-assisted competence	product/performance	AI-assisted practice
on-AI	*about* AI	object of learning	AI-use competence	process/judgment	prompt/orchestrate/review

learning outcomes, rather than primarily as a property of assessment formats, permissions, or lanes. Second, we embed AI-mode within Constructive Alignment, so that decisions about AI use are made across ILOs, ATs, and TLAs rather than within assessment alone. Third, we introduce on-AI as a distinct mode in which AI itself becomes the object of learning. Unlike AI-on, which concerns disciplinary performance with AI, on-AI concerns students' judgment regarding AI itself, including competences such as *prompting*, *orchestrating*, *reviewing*, and *evaluating* AI systems critically.

6.4 AI Tutors, Assistants, and Assessors

A further body of work has examined the use of AI systems as *tutors*, *teaching assistants*, *feedback providers*, and *assessors*. In the context of learning support, Liu et al. describe the use of a course-specific AI chatbot designed to guide students toward solutions rather than provide answers outright [18]. Similarly, Denny et al. found that students valued AI teaching assistants for their *immediacy* and *availability*, but also preferred support that preserved *learner autonomy through scaffolding* rather than *direct solution-giving* [12]. More broadly, Khosravi et al. argue that educational AI systems should be designed as *learning companions* that prioritise *durable learning*, *metacognitive growth*, and *learner agency* over *short-term task performance* [16]. Studies such as these show that AI tutors and assistants can support learning, but only when their design encourages students to engage in the cognitive work required for competence development.

A parallel line of work has examined AI as an assessor or feedback provider. Flodén found that ChatGPT could produce plausible grades on university exams, but also observed *limited grade agreement*, a tendency toward *middle scores*, and *ethical concerns* related to *fairness*, *opacity*, and *accountability* [13]. Usher compared AI, instructor, and peer assessment, finding that AI chatbots could provide *detailed feedback* but sometimes produced *irrelevant* or *contradictory* comments, reinforcing the continuing importance of human judgment [32]. Henderson et al. similarly found that students valued AI-generated feedback for its *accessibility*, *timeliness*, and *clarity*, but considered teacher feedback more *helpful and trustworthy*; they conclude that AI and teacher feedback are complementary rather than interchangeable [14].

These studies demonstrate the *growing range of roles* AI systems can play in teaching, learning, feedback, and assessment. However, they also illustrate why tool-focused accounts are insufficient on their own. The central educational question is not simply *what* an AI tool can do, but what *role* it should play in an aligned learning system. An AI tutor may support an AI-off outcome if it scaffolds students toward autonomous competence *without* completing the task for them. It may support an AI-on outcome if students are learning to work productively *with* AI as part of disciplinary practice. It may support an on-AI outcome if the *interaction itself* becomes an object of reflection and critique. Likewise, AI-assisted assessment is educationally appropriate only if it provides valid evidence of the intended competence and is aligned with the assessment purpose, learning outcome, and surrounding teaching/learning activities. Our framework therefore shifts the focus from the *capabilities* of AI tools to their *role* within an *aligned educational system*.

6.5 Summary

These reviewed works demonstrate growing recognition that AI requires rethinking learning, assessment, and teaching/learning activities in higher education. Existing research has proposed AI literacy frameworks, assessment reforms, and new roles for AI in teaching and assessment. Our contribution complements these developments by extending the underlying educational theory itself. Rather than focusing on AI tools, assessment policies, or AI competences in isolation, we introduce AI-mode as an explicit alignment dimension across *intended learning outcomes*, *assessment tasks*, and *teaching/learning activities* within Constructive Alignment.

7 Conclusion

Generative AI has introduced significant uncertainty into higher education. Responses have ranged from enthusiastic adoption to outright prohibition, and debates often focus on specific tools, policies, or assessments. We have argued that these debates are ultimately manifestations of a deeper issue: *alignment*. Constructive Alignment shifts educational design from intuition and technology-driven reactions toward explicit and principled decisions regarding *intended learning*, including how such learning is *incentivized* and *supported*. In this sense, the challenge posed by AI is ultimately not *technological*, but *educational*.

Our central claim is that AI is not the *cause*; AI is the *catalyst*. By making cognitive outsourcing inexpensive, immediate, and ubiquitous, AI exposes and intensifies longstanding systemic educational misalignments between intended learning outcomes (ILOs), assessment tasks (ATs), and teaching/learning activities (TLAs). At a fundamental level, AI changes means, not ends. The central questions of educational design therefore remain unchanged:

(1) *What do we want students to learn?*
(2) *How do we incentivize it?*
(3) *How do we support it?*

We have argued that, in the age of AI, each of these questions acquires an additional design dimension: AI-mode. Once AI-mode is made explicit, aligned across outcomes, assessment, and activities, and communicated to students, Constructive Alignment provides a principled foundation for educational design in the AI era. Our resulting extension of Constructive Alignment, with the explicit

AI-mode dimension, is summarized in the framework shown in Table 5.

This paper is intended as a stabilizing theory paper for a moment of educational turbulence: a systems response to AI rather than a tool guide, policy statement, or opinion piece. The challenge before us is not to invent entirely new pedagogical theories, but to apply existing ones more rigorously.

In the age of AI, the challenge is not to reinvent educational theory, but to take Constructive Alignment even more seriously.

References

[1] Ibrahim Albluwi, Dennis Bouvier, Claus Brabrand, Michelle Craig, Rodrigo Duran, Chris Hundhausen, Colleen Lewis, Kevin Lin, Andrew Luxton-Reilly, Leo Porter, Karen Reid, David H. Smith IV, Claudia Szabo, Titus Winters, and Daniel Zingaro. 2026. Computing Education when Writing Code is No Longer the Challenge. *Commun. ACM* (2026). To appear.

[2] Ibrahim Albluwi, Dennis Bouvier, Claus Brabrand, Michelle Craig, Rodrigo Duran, Chris Hundhausen, Kevin Lin, Leo Porter, Karen Reid, Andrew Luxton-Reilly, David H. Smith IV, Claudia Szabo, Shubbhi Taneja, Michel Wermelinger, Titus Winters, and Daniel Zingaro. 2026. "Andy's Axe" as a Guiding Principle for our Stance on AI-in-Education (for Computing). In *Proceedings of Dagstuhl Seminar 25311 on Generative AI in Programming Education.*

[3] Lorin W. Anderson and David R. Krathwohl. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* Longman, New York.

[4] John Biggs. 1996. Enhancing teaching through constructive alignment. *Higher Education* 32, 3 (Oct. 1996), 347–364. doi:10.1007/BF00138871

[5] John Biggs, Catherine Tang, and Gregor Kennedy. 2022. *Teaching for quality learning at university 5e.* McGraw-hill education (UK).

[6] John B. Biggs and Kevin F. Collis. 1982. *Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome).* Academic Press, New York.

[7] Benjamin S. Bloom, Max D. Engelhart, Edward J. Furst, Walker H. Hill, and David R. Krathwohl. 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain.* Longmans, Green and Co, New York.

[8] Claus Brabrand. 2026. What Does It Mean to Program in the Age of AI?. In *Proceedings of Dagstuhl Seminar 25311 on Generative AI in Programming Education.*

[9] Claus Brabrand and Jacob Andersen. 2006. Teaching Teaching & Understanding Understanding. https://ttuu.itu.dk/ 19-minute award-winning educational short film.

[10] Claus Brabrand and Sue Sentance. 2026. Use-Modify-Create turned "Upside-Down" by AI: Towards Higher-Level Competences via the Scientific Method. In *Proceedings of the 48th International Conference on Software Engineering (ICSE 2026).*

[11] Adam Bridgeman, Danny Liu, and Ruth Weeks. 2024. Program level assessment design and the two-lane approach. *University of Sydney* (2024).

[12] Paul Denny, Stephen MacNeil, Jaromir Savelka, Leo Porter, and Andrew Luxton-Reilly. 2024. Desirable Characteristics for AI Teaching Assistants in Programming Education. In *Proceedings of the 2024 on Innovation and Technology in Computer Science Education V. 1* (Milan, Italy) *(ITiCSE 2024).* Association for Computing Machinery, New York, NY, USA, 408–414. doi:10.1145/3649217.3653574

[13] Jonas Flodén. 2025. Grading exams using large language models: A comparison between human and AI grading of exams in higher education using ChatGPT. *British Educational Research Journal* 51, 1 (2025), 201–224. doi:10.1002/berj.4069

[14] Michael Henderson, Margaret Bearman, Jennifer Chung, Tim Fawns, Simon Buckingham Shum, Kelly E. Matthews, and Jimena de Mello Heredia. 2025. Comparing Generative AI and teacher feedback: student perceptions of usefulness and trustworthiness. *Assessment & Evaluation in Higher Education* 0, 0 (2025), 1–16. doi:10.1080/02602938.2025.2502582

[15] Peter Kandlbinder. 2014. Constructive alignment in university teaching. *HERDSA News* 36, 3 (2014), 5–6. https://search.informit.org/doi/10.3316/informit.150744867894569

[16] Hassan Khosravi, Dragan Gasevic, Shazia Sadiq, Lixiang Yan, Jason Lodge, Jason Tangen, Paul Denny, Kristen DiCerbo, Simon Buckingham Shum, and Ryan S. Baker. 2026. Building AI Companions that Prioritise Learning over Performance. arXiv:2605.04816 [cs.HC] https://arxiv.org/abs/2605.04816

[17] Rongxin Liu, Benjamin Xu, Christopher Perez, Julianna Zhao, Yuliia Zhukovets, and David J. Malan. 2025. Assessment in CS50 with AI: Leveraging Generative Artificial Intelligence for Personalized Student Evaluation. In *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 2* (Pittsburgh, PA, USA) *(SIGCSETS 2025).* Association for Computing Machinery, New York, NY, USA, 1735. doi:10.1145/3641555.3705061

[18] Rongxin Liu, Carter Zenke, Charlie Liu, Andrew Holmes, Patrick Thornton, and David J. Malan. 2024. Teaching CS50 with AI: Leveraging Generative Artificial Intelligence in Computer Science Education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1* (Portland, OR, USA) *(SIGCSE 2024).* Association for Computing Machinery, New York, NY, USA, 750–756. doi:10.1145/3626252.3630938

[19] Jason Lodge, Sarah Howard, Margaret Lynn Bearman, Phillip Dawson, Shirley Agostinho, Simon J. Buckingham Shum, Chris Deneen, Cath Ellis, Tim Fawns, Helen Gniel, Rowena Harper, Michael Henderson, Danny Yen Ting Liu, Lina Markauskaite, Jan McLean, Carlo Perrotta, Lambert Schuwirth, Christine Slade, and Associates. 2023. *Assessment Reform for The Age of Artificial Intelligence.* Tertiary Education Quality and Standards Agency.

[20] Jason M. Lodge and Leslie Loble. 2026. *Artificial Intelligence, Cognitive Offloading and Implications for Education.* Technical Report. University of Technology Sydney. doi:10.71741/4pyxmbnjaq.31302475

[21] Colin Loughlin, Simon Lygo-Baker, and Åsa Lindberg-Sand. 2021. Reclaiming constructive alignment. *European Journal of Higher Education* 11, 2 (2021), 119–136. doi:10.1080/21568235.2020.1816197

[22] Fengchun Miao, Kelly Shiohira, et al. 2024. *AI competency framework for students.* UNESCO Publishing.

[23] Davy Tsz Kit Ng, Jac Ka Lok Leung, Samuel Kai Wah Chu, and Maggie Shen Qiao. 2021. Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence* 2 (2021), 100041. doi:10.1016/j.caeai.2021.100041

[24] OECD and European Union. 2026. *Empowering Learners for the Age of AI: An AI Literacy Framework for Primary and Secondary Education.* Technical Report. OECD Publishing, Paris. doi:10.1787/65cd27d4-en

[25] Mike Perkins, Leon Furze, Jasper Roe, and Jason MacVaugh. 2024. The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment. *Journal of University Teaching and Learning Practice* 21, 6 (2024), 49–66. https://search.informit.org/doi/10.3316/informit.T2024092900003300954126858

[26] Mike Perkins, Jasper Roe, and Leon Furze. 2025. Reimagining the Artificial Intelligence Assessment Scale (AIAS): A Refined Framework for Educational Assessment. *Journal of University Teaching and Learning Practice* 22, 7 (2025), 1–26. doi:10.53761/rrm4y757

[27] Victor-Alexandru Pădurean, Paul Denny, Alkis Gotovos, and Adish Singla. 2025. Prompt Programming: A Platform for Dialogue-based Computational Problem Solving with Generative AI Models. In *Proceedings of the 30th ACM Conference on Innovation and Technology in Computer Science Education V. 1* (Nijmegen, Netherlands) *(ITiCSE 2025).* Association for Computing Machinery, New York, NY, USA, 458–464. doi:10.1145/3724363.3729094

[28] Yunhan Qiao, Md Istiak Hossain Shihab, and Christopher Hundhausen. 2026. A Systematic Literature Review of the Use of GenAI Assistants for Code Comprehension: Implications for Computing Education Research and Practice. *ACM Trans. Comput. Educ.* 26, 2, Article 24 (Feb. 2026), 33 pages. doi:10.1145/3785366

[29] Christian Stamov Roßnagel, Noleine Fitzallen, and Katrin Lo Baido. 2021. Constructive alignment and the learning experience: relationships with student motivation and perceived learning demands. *Higher Education Research & Development* 40, 4 (2021), 838–851. doi:10.1080/07294360.2020.1787956

[30] Stephanie Santschi. 2026. Think tasks, not tools: why AI policy in higher education needs constructive alignment. *Teaching in Higher Education* 0, 0 (2026), 1–11. doi:10.1080/13562517.2026.2671391

[31] Hung-Yue Suen and Kuo-En Hung. 2023. Building trust in automatic video interviews using various AI interfaces: Tangibility, immediacy, and transparency. *Computers in Human Behavior* 143 (2023), 107713. doi:10.1016/j.chb.2023.107713

[32] Maya Usher. 2025. Generative AI vs. instructor vs. peer assessments: a comparison of grading and feedback in higher education. *Assessment & Evaluation in Higher Education* 50, 6 (2025), 912–927. doi:10.1080/02602938.2025.2487495

[33] Lev S. Vygotsky. 1978. *Mind in Society: The Development of Higher Psychological Processes.* Harvard University Press, Cambridge, MA.

[34] Ying Xiong and Joon Kyoung Kim. 2025. Who wants to be hired by AI? How message frames and AI transparency impact individuals' attitudes and behaviors toward companies using AI in hiring. *Computers in Human Behavior: Artificial Humans* 3 (2025), 100120. doi:10.1016/j.chbah.2025.100120